

침수피해와 침수취약요인 관계 연구

: 피해액과 인명피해를 중심으로*

Relationship between Flood Damage and Flood Vulnerability Focusing on Property Damage and Human Casualties

박한나** · 송재민***

Park, Hanna · Song, Jaemin

Abstract

This study advocates developing efficient disaster management strategies and recognizing potential risk areas by studying areas historically affected by severe flood damage. Prior research in Korea has predominantly concentrated on property damage as the main signifier of flood impact, but human casualties are equally significant. Machine learning methods were used in this study to split flood damage into property damage and human casualties. We also explore the connection between damage characteristics and factors contributing to flood vulnerability. The Gradient Boosting and Random Forest models were identified as the most appropriate for forecasting property damage and human casualties, respectively. This study reveals that various elements influence property damage and human casualties. Physical elements and high rainfall exposure cause property damage, while demographic attributes and the drainage capability of impacted regions influence human casualties. Thus, this research underscores the need to consider property damage and human casualties when determining relevant policies. By differentiating between property damage and human casualties and examining the nonlinear connection between flood damage and key vulnerability features, this study aids urban planning initiatives to foster flood-resistant communities.

주제어 침수피해, 피해액, 인명피해, 침수취약성 요인, 기계학습모형

Keywords Flood Damage, Property Damage, Human Casualties, Flood Vulnerability Indicator, Machine Learning Model

1. 서론

기후변화는 현재 지구의 지속가능성을 위협하는 가장 큰 리스크이다(IPCC, 2014). 기후변화로 인해 폭우, 폭염, 태풍 등의 이상 기온 현상이 나타나는 발생 빈도가 빈번해지고 있으며, 피해 규모도 대형화되고 있다(IPCC, 2014; Jongman et al., 2018). 특히 우리나라는 자연재해 중 침수피해로 발생하는 피해액이 연

간 전체 피해액의 89% 이상을 차지함에 따라 침수피해 및 기후변화 대응 정책과 적절한 도시계획의 필요성이 증가하고 있다.

기후변화 대응 계획 및 정책 수립을 위해서는 피해 지역에 대한 식별과 양상을 파악하는 것이 우선적으로 필요하다. 우리나라의 경우 인적 피해에 대한 이해도 중요하지만 대다수의 침수피해 관련 선행연구에서 자연재해 피해액이 주로 활용되어왔다(최충익, 2003; 최열·서만훈, 2013; 서만훈 외, 2016; 심재현 외, 2012;

* 이 연구는 서울대학교 환경계획연구소에서 지원되는 연구비에 의하여 수행되었음.

** Ph.D. Candidate, Graduate School of Environmental Studies, Seoul National University (First Author: hnpark@snu.ac.kr)

*** Associate Professor, Graduate School of Environmental Studies, Seoul National University (Corresponding Author: jaemins@snu.ac.kr)

Jeong and Yoon, 2018). 자연재해 피해액은 침수, 산사태, 폭설, 강풍 등 자연재해로 발생하는 건물, 선박, 도로, 사유재산 등에 발생하는 물리적인 피해를 금전적인 가치로 산정한 것이다. 하지만, 자연재해 피해를 고려할 때 자연재해로 인해 사망 및 부상을 입은 인적 피해에 대해 고려하는 것이 필요하다. 왜냐하면 어떠한 피해를 고려하는지에 따라 취약요인이 다를 수 있기 때문이다. 예를 들어, 동일한 수준의 재해 노출에도 건물 밀도가 높고, 토지의 공시지가가 큰 대도시에서 재해 피해액이 높게 산정되는 경향이 있다(최열·서만훈, 2013; Yoon, 2012). 반면, 인적 피해의 경우 지역의 재해대응 역량이 낮은 비도시지역에서 더 높게 나타나는 경향이 있다(Tellman et al., 2020; Jamshed et al., 2020). 이와 같이, 침수피해를 물리적인 피해로 보는지 인적 피해로 측정하는지에 따라 파악되는 피해 양상이 매우 달라질 수 있다.

하지만 자연재해로 발생하는 인명피해는 사회 재난 및 안전사고에 비해 발생 빈도가 낮은 특성으로 인명피해에 대한 고려 중요성이 상대적으로 낮게 인식되어왔다(조현우·최승용, 2018; 홍승진 외, 2018). 더욱이 자연재해로 발생하는 사망, 실종 및 부상자를 통합적으로 고려할 때 상호 간 명확한 환산 기준이 없어, 사망자만 인명피해로 보기도 하고(Tellman et al., 2020; Bakkenen et al., 2017), 사망자와 부상자를 가중치 없이 단순히 산술적으로 합산하여 인명피해로 정의하기도 한다(Zahran et al., 2008; 이종석·최현일, 2016; 김모정·김광섭, 2018). 이와 같은 이유로 특히 우리나라의 경우 침수 리스크 모델 및 기존 선행연구에서 인명피해 지표가 널리 활용되지 않았다. 하지만, 방재정책에서 물적 피해를 줄이는 것도 중요하지만 동시에 인명피해 최소가 핵심 목표가 되어야 한다는 점을 고려할 때, 폭우에 대한 회복탄력적 도시를 구축하기 위해서는 피해액과 인명피해를 구분하여 침수피해 양상과 영향요인을 분석하는 심도 있는 연구가 필요하다.

이와 같은 배경에서, 본 연구는 침수피해를 피해액과 인명피해로 구분하고, 피해 특성과 침수취약성 간 관계를 규명하는 것을 목적으로 한다. 구체적으로는 첫째, 침수피해의 주요 지표인 피해액과 인명피해 간의 차이를 파악하고, 둘째, 피해액과 인명피해 발생에 기여하는 주요한 침수 취약요인의 차이와 비선형적 기여도에 대한 이해를 넓히고자 한다. 이를 위해 우선, 선행연구 검토를 통해 피해액과 인명피해에 영향을 미치는 침수취약성 지표를 식별하고, 기계학습 모델을 이용하여 침수피해 양상과 침수취약요인 간 비선형적인 관계를 밝히고자 한다. 국내의 기존 자연재해 관련 연구에서는 일반적으로 종속변수와 독립변수 간 선형적인 관계를 가정하는 선형회귀 모델을 활용하거나 종속변수에 0의 분포가 많은 특성을 고려한 영과잉 음이항 회귀모형을 활용하여 침수피해 영향요인을 식별하였다. 하지만 재해피해는 일상적으로 발생하는 것이 아닌, 특정 시점에 극단적인 피해가 발생

하는 특성이 있다. 이에 도시의 물리적 환경과 인구·사회적 특성이 침수피해에 미치는 영향을 선형관계만으로 진단하기에는 한계가 있다. 이러한 흐름에서 최근 홍수 등 자연재해 분야에서는 피해를 예측하고 재해 취약성을 평가하기 위해 기계학습 모형 사용이 늘어나고 있다(Ma et al., 2021; Parvin et al., 2022; Pham et al., 2021; Ha and Kang, 2022). 본 연구는 기존 피해액 위주의 침수피해 연구에서 인명피해로 침수피해 양상을 다각화하고, 침수피해 양상과 침수취약요인 간 비선형적인 관계를 포착함으로써 기존 침수피해 연구를 보완하고 있으며, 물적 피해와 인적 피해에 영향을 미치는 주요한 침수취약 특성을 규명하여 효과적인 방재정책에 시사점을 제시한다는 데에 의의가 있다.

II. 선행연구 고찰

1. 침수취약성 지표

자연재해 취약성은 일반적으로 기후 노출, 민감도, 적응역량의 함수로 정의되며, 이는 침수피해를 산정하고 잠재적인 자연재해 취약성을 검토하기 위해 널리 활용되고 있다(IPCC, 2014; Bigi et al., 2021; Salazar-Briones et al., 2020; Sayers et al., 2018; Chakraborty et al., 2020). 침수취약성 구성요인 중 기후 노출은 재해를 유발하는 기후적 요소에 의한 영향, 민감도는 동일한 재해 노출에서 더 큰 피해를 입을 수 있는 물리적, 사회·경제적 취약요인을 의미하며, 적응 능력은 피해를 저감하고 회복력을 높이기 위한 요인으로 정의된다. 침수 취약성 지표는 물리적 요인뿐만 아니라 사회경제적 요인까지 포괄하고 있는데(IPCC, 2014; Rufat et al., 2019), 이는 지형적 및 물리적, 기술적 원인만 고려하여 침수취약성을 평가했던 과거의 방법과 차별되는 점이다(Nasiri et al., 2016). 최근 침수취약성은 물리적, 환경적, 제도적, 사회적, 경제적 요인 등 다차원으로 침수취약 요인을 고려하여, 통합지표로 평가되는 것이 적절하다고 보고 있다(Salas and Yepes, 2018).

선행연구에서 제시하고 있는 구체적인 침수취약성 세부지표는 다음 <표 1>과 같이 구성된다. Bigi et al.(2021)는 도시지역 침수취약성 지표에 대한 선행연구를 바탕으로 공통적으로 자주 사용되는 하위 지표를 제시하고 있다. 인구밀도, 소득, 나이 등 취약 인구 특징과 저지대와 급경사지, 노후주택 등 물리적 취약성, 그리고 녹지면적, 하수관거 등 물리적 대응력이 이에 해당한다. Sayers et al.(2018)은 영국을 대상으로 근린주구 단위의 침수취약성을 개발하여 커뮤니티의 재해대응력을 고려하였다. 커뮤니티 단위에서 침수취약성을 도출할 때 거주 기간 등을 활용하여 사회적 네트워크를 고려하고, 영어 사용 능력으로 정보의 이용성을 고려하며, 방문요양 시설 및 보건 의료시설의 서비스 이용성을 고려하는 것이 특징적이다. 다수의 연구에서는 침수 취약성 요인

Table 1. Indicators related to Flood Vulnerability Index in the previous research

Component		Indicator	References	
Exposure Index (EI)	Climate factors	Average of maximum daily rainfall	Pathak et al. (2020), Chuang et al. (2020), Kang and Lee (2012), Kim, J. et al. (2013), Park and Song (2014)	
		Number of days with more than 80 mm rainfall	Pathak et al. (2020), Chuang et al. (2020), Munyai et al. (2019), Kang and Lee (2012), Park and Song (2014)	
	Physical factors	Area within 100 m from rivers	Pathak et al. (2020), Salazar-Briones et al. (2020), Kang and Lee (2012), Kim, S.J. et al. (2013), Seo and Han (2019)	
		Density of steep slopes	Salazar-Briones et al. (2020), Munyai et al. (2019), Kang and Lee (2012), Kim, S.J. et al. (2013), Seo and Han (2019), Shin et al. (2019)	
		Ratio of low-lying land below an average altitude	Lee et al. (2020), Kim, J. et al. (2013), Kim, S.J. et al. (2013), Seo and Han (2019)	
		Ratio of impermeable area	Chuang et al. (2020), Bigi et al. (2021), Kang and Lee (2012), Lee et al. (2020), Kim, J. et al. (2013), Shin et al. (2019)	
Sensitivity Index (SI)	Population factors	Ratio of old single houses	Bigi et al. (2021), Lee et al. (2020), Seo and Han (2019)	
		Population density	Salazar-Briones et al. (2020), Tascon-Gonzalez et al. (2020), Chuang et al. (2020), Munyai et al. (2019), Kim, J. et al. (2013), Shin et al. (2019)	
		Ratio of older single households	Kim et al. (2018), Park et al. (2020), Seo and Han (2019)	
	Economical factors	Percentage of people aged over 65 or under 5 years old	Salazar-Briones et al. (2020), Tate et al. (2021), Seo and Han (2019)	
		Percentage of economically vulnerable group	Pathak et al. (2020), Tascon-Gonzalez et al. (2020)	
		Unemployment rate	Bigi et al. (2021), Kim et al. (2018), Tate et al. (2021), Munyai et al. (2019)	
		Property density	Kim et al. (2018), NDMI (2011), Park et al. (2020), Lee et al. (2020), Shin et al. (2019)	
		Physical factors	Ratio of embankment area	Kim et al. (2018), Munyai et al. (2019), NDMI (2011)
			Ratio of green area	Bigi et al. (2021), Kang and Lee (2012), Kim, J. et al. (2013), Park and Song (2014)
Length of storm sewer per area	Bigi et al. (2021), Kim, J. et al. (2013), Park and Song (2014), Shin et al. (2019)			
Adaptive Capacity Index (ACI)	Population factors	Number of medical and health workers	Tascon-Gonzalez et al. (2020), Kim et al. (2018), Park and Song (2014)	
		Number of firefighters and police officers	Tascon-Gonzalez et al. (2020), Munyai et al. (2019), Kim, J. et al. (2013)	
		Ratio of the population who are under elementary education	Jeong and Yoon (2018)	
	Economical factors	Level of Financial independence	Shin et al. (2019), Kim, J. et al. (2013), Park and Song (2014)	
		Ratio of insurance subscribers for flood damage	Pathak et al. (2020), Munyai et al. (2019)	

이 다양하기 때문에, 주성분 분석을 활용하여 통계적으로 유의미한 요인을 저차원으로 축소하여 도출하기도 한다(Chakraborty et al., 2020; Yoon, 2012).

한편, 침수취약성 지표를 통합적으로 산출하기 위해서 다양한 변수들에 가중치를 반영하여 하나의 통합지표로 활용한다. 가중치 산정 방법에는 산술평균으로 동일한 가중치를 반영하기도 하고(Salazar-Briones et al., 2020), 분석적 계층화 과정(AHP, Analytical Hierarchical Process)을 따라 전문가 설문조사를

활용하기도 하고, 각 정보 속성의 크기분포를 이용하여 응집도를 결정하는 엔트로피 방법(Entropy)을 활용하기도 한다(Tascon-Gonzales et al., 2020; 박하나·송재민, 2014).

2. 침수취약성 요인이 침수피해에 미치는 영향

1) 피해액

다수의 선행연구에서는 침수피해를 산정하는 지표로 자연재해

피해액을 주로 사용한다(Tellman et al., 2020; Jeong and Yoon, 2018; 최열·서만훈, 2013). 자연재해 피해액을 활용한 다수의 실증연구에서는 도시지역 특성 요인이 침수피해를 가중시키는 주요한 요인으로 도출되었다(최충익, 2003; 최열·서만훈, 2013; 서만훈 외, 2016). 이는 인구와 예산, 건물 밀도가 높은 도시지역에서 침수피해액이 더 크게 나타나는 특성으로 인해 도시지역의 토지이용 및 사회·경제적, 건설환경, 인구적 특성 요인이 침수피해액을 가중시키는 영향으로 나타난 것이다. 하지만 다른 연구에서는 낮은 주거 밀도 특성을 갖는 비도시지역의 토지이용 특성을 침수피해액을 가중시키는 요인으로 보기도 한다(Tellman et al., 2020; Jeong and Yoon, 2018). 도시지역 특성으로 대표되는 밀도 높은 주택단지(Compact community)는 오히려 녹지로 구성된 오픈스페이스의 면적을 확대함으로써 빗물 투수율을 높이고, 빗물의 표면유출을 저감하여 침수피해를 줄이기 때문이다(Jeong and Yoon, 2018). 따라서 자연환경 및 건설환경 요인 중 높은 주거 밀도와 넓은 녹지공간은 재해피해액을 감소시키고, 높은 경사도와 강수량은 재해피해액을 증가시키는 것으로 나타났다(Jeong and Yoon, 2018). 높은 강수량은 강한 재해 노출 요인에 해당하며, 평균 경사도가 높은 지역은 우수유출을 가중시켜 침수 발생 확률이 증가하기 때문에 침수피해액이 커지는 요인으로 본다(장미홍·박창열, 2019; Jeong and Yoon, 2018).

한편, 자연재해 피해액에 영향을 미치는 또 다른 요인으로는 지역의 인구 특성이 있다. 특히, 기존 선행연구에서는 경제적 취약성을 나타내는 저소득층과 저교육층, 유색인종과 소수민족 요인이 재해피해액을 가중시키는 것으로 나타났다(Tellman et al., 2020; Jeong and Yoon, 2018). 이는 경제적 취약계층은 상대적으로 침수 위험이 높은 지역에서 이주하지 못하고 거주함으로써 홍수 위험에 높게 노출되며, 제방시설과 같은 홍수 저감 인프라 투자로부터 소외되고, 홍수보험과 같은 금융적 지원도 매우 저조하기 때문이다(Bakkensen et al., 2017; Rufat et al., 2019; Tellman et al., 2020).

마지막으로, 경제적 요인 중 지역 및 개인의 경제 수준이 재해 피해액과 관련이 있는 것으로 나타났다. 지역사회의 빈곤은 비상 물품의 보관과 홍수보험의 가입을 어렵게 하고, 재해피해로부터 회복력을 약화시키는 원인이 되기 때문에 지역의 경제력이 낮을수록 재해피해액이 높아지는 결과를 낳는다(Jeong and Yoon, 2018). 이는 가난한 유색인종이 홍수 관련 재산손실이 높은 경향을 보이는 것과 같은 맥락이다(Tellman et al., 2020). 한편, 주택가격이 낮은 지역은 열악한 주거환경과 정부의 인프라 투자로부터 소외됨으로써 침수피해액이 높게 나타났다(Tellman et al., 2020).

2) 인명피해(사상자 및 부상자)

해의 침수와 관련한 인명피해 연구는 주로 사회적 취약성과 인

명피해 간의 관계를 규명한다(Zahran et al., 2008; Bakkensen et al., 2017; Tellman et al., 2020). 논의의 흐름은 미국을 중심으로 2005년 Katrina 태풍 피해 이후, 환경정의(Environmental Justice) 대상이 자연재해까지 확장됨에 따라, 지역의 사회적 취약성이 침수피해에 영향이 있는지 밝히는 데 기반한다(Maantay and Maroko, 2009; Cutter and Finch, 2008; Tate et al., 2021). 선행연구에 따르면 침수로 인한 인명피해는 소수민족이나 유색인종, 경제적 취약계층, 노인 및 영유아 등 사회적·경제적 취약계층이 높게 거주하는 지역에서 높게 나타나는 것으로 밝혀졌다(Zahran et al., 2008; Bakkensen et al., 2017; Tellman et al., 2020). 이는 인명피해가 실수나 인재에 의해 발생하는 것보다 지역의 취약성 요인이 인명피해의 발생에 유의미한 영향을 미치는 것을 의미한다. 사회적 취약계층의 거주지역에 침수로 인한 인명피해 발생이 높게 발생하는 이유는 상습침수 발생지역의 낮은 자가로 저소득계층이 거주지를 선택하는 요인과 지방재정의 방재 및 구급 대응력 강화에 투자가 미흡한 요인, 신체적 취약계층의 대피 능력이 부족한 요인 등 지역의 일반적인 특성과 관련된다(Bakkensen et al., 2017; Rufat et al., 2019; Tellman et al., 2020).

또한 침수 인명피해는 재해 노출 특성과 물리적·경제적 대응력 요인도 영향을 미치는 것으로 나타났다. 돌발홍수나 해안홍수와 같이 급작스럽게 발생하는 재해 노출 상황에서 급류에 휩쓸리거나 건물·차량에 갇혀 발생하는 인명피해가 많았다(Ashley and Ashley, 2008; Jonkman and Kelman, 2005; Jonkman, 2014). 물리적 대응력 요인으로 댐의 수와 경제적 대응력 요인으로 홍수 피해보험 가입률이 높을수록 침수 인명피해가 감소하는 영향을 보인다(Zahran et al., 2008). 이는 댐의 수위조절로 강물의 범람을 예방하고, 홍수피해보험은 침수로 인한 경제적 손실에 대비할 수 있는 지역의 경제적 역량을 반영하기 때문이다.

국내 침수와 관련한 인명피해 연구는 산정된 침수취약성 지표를 검증하기 위한 피해지표로 피해액과 더불어 인명피해 및 침수 면적이 활용되거나(김묘정·김광섭, 2018; 이종석·최현일, 2016), 피해지표별로 침수취약 지역을 선정하여 인구특성을 분석하기 위해 인명피해 지표가 활용된다(Lee et al., 2020). 이때 인명피해는 사망자 수만 반영하기도 하고(조현우·최승용, 2018), 사망·실종·부상자의 합산으로 정의하기도 한다(이종석·최현일, 2016; 김묘정·김광섭, 2018). 이는 자연재해 분야에서 인명피해에 대한 가치 환산 기준이 정립되지 않은 한계로 대두되는데, 사회재난의 경우 사망자와 부상자에 대한 가치 기준이 철도사고 및 건설사업장 산업재해 부문에 대해 정의¹⁾되어 있어 이를 활용하기도 한다(이관형 외, 2016; 조심정, 2018; 조심정·최승용, 2018). 사회재난에 비해 자연재해로 인한 인명피해는 발생빈도가 낮고, 규모도 작으며, 특히, 침수로 인한 부상자는 집계하기 어려운 한계가 있어 일반적으로 사망자보다 부상자의 수가 적게 집계

되는 특성이 있다(Doocy et al., 2013; 홍승진 외, 2018). 이러한 발생특성으로 자연재해 인명피해는 상대적 중요도가 낮게 인식되어왔고, 사회재난과 다르게 등가 사망 환산 기준이 부재하다(홍승진 외, 2018; 조현우·최승용, 2018; 국토교통부, 2018). 인명피해 가치환산의 어려움은 기존 침수피해 연구에서 인명피해를 주요 피해지표로 고려하기에 한계로 작용하였다.

한편, 인명피해는 피해가 발생하지 않은 0의 값이 많은 자료의 분포특성에 따라, 선행연구에서는 영과잉 음이항 회귀모형(Zero-inflated Negative Binomial Regression)을 활용하여 사회적 취약성과 침수 인명피해의 영향 관계를 분석한다(Tellman et al., 2020; Zahran et al., 2008). 영과잉 음이항 회귀모형(Zero-inflated Negative Binomial Regression)은 사건이 발생하지 않은 집단과 사건이 발생한 집단을 구분하여 0이 발생할 확률(사건이 발생하지 않을 확률)을 구하고, 음이항 회귀모형으로 0이 아닌 나머지 관측치의 발생확률(사건발생 확률)을 추정하는 모형이다. 따라서 한두 건의 작은 인명피해 관측치도 모형에서 0의 발생확률과 가산자료 발생확률을 구분하여 추정하고, 가산자료의 발생확률 추정은 포아송분포의 이분산성을 허용하는 음이항 회귀를 이용하기 때문에 통계적 유의성을 확보할 수 있다.

3. 기존 연구와의 차별성

인명피해와 피해액에 영향을 미치는 침수취약성 요인은 <표 2>와 같이 공통점과 차이점이 존재한다. 해외 선행연구에서 피해액과 인명피해를 구분하여 침수피해와 관계를 검토해온 것에 비해 우리나라에서는 침수피해 지표로 인명피해를 고려한 연구가 매우 부족하다. 이는 침수로 발생하는 피해액은 발생 빈도도 많고, 규모도 큰 반면, 침수 인명피해는 피해액에 비해 상대적으로 발생 빈도도 낮고, 규모도 작기 때문에 침수피해 지표로 인명피해의 중요도가 낮게 인식되어온 것이다(홍승진 외, 2018). 더욱이 인명피해의 낮은 발생 빈도와 이벤트성 데이터 분포로 선행회귀 모형의 통계적 가정에 적합하지 않은 한계가 있어 기존 선행연구에서 인명피해는 주된 관심사가 아니었다. 하지만 해외 연구에서는 인명피해 발생 특성을 고려한 통계적 모형으로 검증하고 있으며, 해외 선행연구 결과에 따르면, 피해액과 인명피해에 영향을 미치는 침수취약성 요인은 다르게 나타났다(Tellman et al., 2020; Zahran et al., 2008). 이는 피해지표에 따라 침수피해의 발생 기제가 다름을 시사한다.

본 연구는 국내를 대상으로 피해액과 인명피해 발생에 영향을 미치는 침수 취약성 요인을 밝힘으로써 기존 재해대응정책에서 간과하고 있던 침수피해 지표에 따른 발생 기제의 차이를 조명하는데 의의가 있다. 방법론 차원에서는 자료의 사전적 가정이 없는 비모수적 통계기법인 기계학습 모형을 활용하여 침수피해와 침수취약 요인 간 비선형적인 관계를 포착함으로써, 전통적인 통계

모형을 활용하여 변수 간 선형관계를 포착하는 기존 연구와 차별점을 가진다. 트리 기반 기계학습 모형은 비모수적 접근으로 모집단의 가정을 필요로 하지 않고, 예측치와 실제값의 차이를 최소화 하는 데이터 특성값을 찾는 규칙에 의해 학습하고, 예측한다. 본 연구에서 사용한 앙상블 학습은 설명변수의 다양한 조합으로 표본을 추출하고, 각각의 트리 기반 기계학습 모형으로 예측한 값들을 평균값 또는 최빈값으로 통합하기 때문에 분산을 줄이며, 정확도를 높이는 효과가 있다(최창현 외, 2018; Chen et al., 2020). 이러한 연산으로 기계학습 모형은 이상치나 노이즈에 강건한 특성이 있기 때문에 침수피해와 같이 극단적인 이벤트성 데이터에 영향요인을 설명하기에 전통적인 회귀모형보다 기계학습 모형의 설명력이 더 높다고 알려져 있다(Tehrany et al., 2014; Yoon and Jeong, 2016; Rahmati and Pourghasemi, 2017). 이에 본 연구에서 피해액과 인명피해에 영향을 미치는 침수취약 요인을 트리 기반 기계학습 모형으로 검증함으로써 기존 통계 모형에서 데이터 분포가 가정을 위배하여 발생할 수 있는 모형 추정의 오류를 개선하고, 추정모형의 정확도를 높인다.

III. 연구방법

1. 분석의 틀

본 연구의 목적은 침수피해를 피해액과 인명피해로 구분하여 침수피해 양상을 분석하고, 침수피해별 침수취약 요인 간 비선형적인 관계를 분석하는 것이다. 본 연구는 <그림 1>과 같은 연구 흐름을 통해 진행되며, 다음의 세 가지 연구 질문에 답하고자 하였다. 구체적인 연구 질문과 방법은 다음과 같다.

첫째, 침수피해를 나타내는 피해액과 인명피해 간에는 어떠한

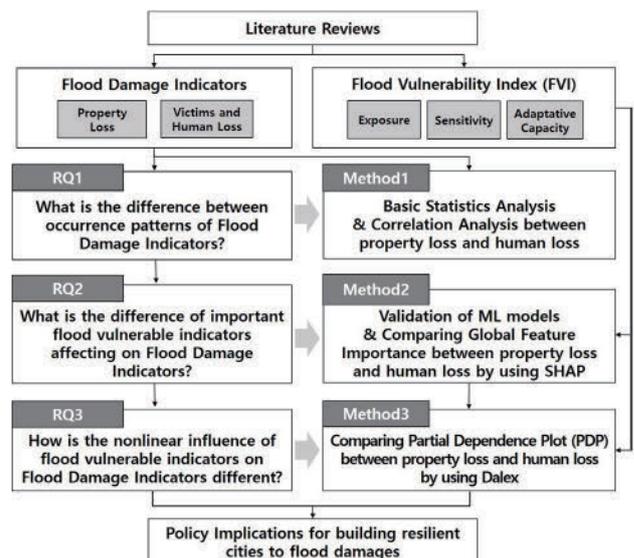


Figure 1. Research flow of the study

Table 2. Causes of property damage and human casualties due to flooding

Indicator	Property damage	Flood human casualties	References
Natural environmental characteristics	Flood type		○
	Precipitation	○	○
	Average slope	○	
	River	○	
	Low lands	○	
Population characteristics	Age (children, elderly)	○	○
	Sex (female)	○	○
	Race (color)	○	○
	Population change		○
	Population density	○	○
	Education level	○	
	Total population	○	○
Land use characteristics	Rural area	○	○
	Urban area	○	
	Green and park area	○	
	Impermeable area	○	
	Wetland, bare ground	○	
Construction environmental characteristics	Housing density	○	
	Urbanized area	○	
	Building floor area	○	
	Sewage pipe	○	
	Vulnerable residential type	○	
	Old buildings	○	
	Road density	○	
Socioeconomic characteristics	Income	○	○
	Flood damage insurance		○
	Housing prices	○	
	Unemployment rate	○	
	Manufacturing ratio	○	
	The size of the city budget	○	
	GDP per population		○
Disaster defense characteristics	Developing countries		○
	The area of the embankment	○	
	Number of dams		○
	Reservoirs, reservoir area	○	

▪ Property damage: Jeong and Yoon (2018), Choi (2003)
 ▪ Human casualties: Zahran et al. (2008), Doocy et al. (2013)

▪ Property damage: Yoon (2012), Choi (2003), Shim et al. (2012), Seo et al. (2016), Choi and Seo (2013), Tellman et al. (2020)
 ▪ Human casualties: Lee et al. (2020), Zahran et al. (2008), Tellman et al. (2020), Doocy et al. (2013)

▪ Property damage: Yoon (2012), Jeong and Yoon (2018), Choi (2003), Shim et al. (2012), Seo et al. (2016), Choi and Seo (2013), Tellman et al. (2020)
 ▪ Human casualties: Tellman et al. (2020)

▪ Property damage: Yoon (2012), Jeong and Yoon (2018), Seo et al. (2016), Choi and Seo (2013), Tellman et al. (2020)

▪ Property damage: Yoon (2012), Jeong and Yoon (2018), Seo et al. (2016), Tellman et al. (2020)
 ▪ Human casualties: Zahran et al. (2008), Tellman et al. (2020), Jonkman (2014), Doocy et al. (2013)

▪ Property damage: Shim et al. (2012)
 ▪ Human casualties: Zahran et al. (2008)

차이가 존재하는가? 이를 위해 침수피해 지표의 기초통계분석과 자료의 분포 특성을 분석하고, 피해액과 인명피해 간 상관관계 분석 및 공간적 분포 패턴 분석을 통하여 두 개의 침수피해 지표 간 관련성을 분석하였다. 둘째, 피해액과 인명피해 발생에 기여하는 주요한 침수취약 요인의 차이는 무엇인가? 침수피해에 영향

을 미치는 침수취약 요인을 분석하기 위해 본 연구에서는 전통 회귀모형과 다양한 종류의 머신러닝 모형을 비교하고, R-Squared 값과 표준절대오차(MAE, Mean Absolute Error), 평균제곱근 오차(RMSE, Root Mean Square Error)를 비교하여 가장 우수한 모형을 선정한다. 선정된 모형에서 피해액과 인명피해의 예측

력에 영향을 미치는 침수취약성 요인의 중요도를 각각 도출하고, 이에 대한 차이를 비교한다. 셋째, 피해액과 인명피해에 대한 침수취약성 요인의 비선형적 영향력은 어떻게 다른가? Partial Dependence Plot(PDP)을 활용하여 기계학습 모형에서 침수피해의 예측력에 중요하게 도출된 침수취약성 요인의 한계효과(marginal effect)를 해석한다. PDP는 관심변수를 제외한 나머지 변수들은 평균값으로 고정시킨 상태에서, 관심 있는 변수의 값을 변화시키며 예측값을 계산한 뒤, 예측값들의 평균을 2차원 그래프로 나타낸 것으로, 종속변수와 설명변수의 선형관계뿐만 아니라 비선형관계도 설명할 수 있는 방법이다.

2. 변수의 구성

본 연구는 대한민국 229개 시·군·구를 공간적 범위로 하고, 2019년을 시간적 기준으로 한다. 침수피해는 측정 시기별 편차가 발생하기 때문에 2010년부터 2019년까지 10년간 총합으로 한다. 종속변수는 시·군·구별 총인구수로 나누어 단위값이 보정된 피해액과 인명피해로 하며, 행정안전부에서 해마다 발간하는 재해연보의 자료를 활용한다. 피해액과 인명피해 데이터는 재해연보에 제시되는 229개 시·군·구의 자연재해 피해액과 사망·실종·부상자 수를 합산하여 사용하였으며(Zahran et al., 2008; 이종석·최현일, 2016; 김묘정·김광섭, 2018),²⁾ 2010년부터 2019년까지 10년간 평균값으로 반영한다. 이때 1인당 피해액과 백만인당 인명피해로 표준화한다.

설명변수로 활용되는 침수취약성 구성요인은 IPCC(2007)의 Vulnerability 개념을 참고하여, 재해노출 지수(EI, Exposure Index), 민감도 지수(SI, Sensitivity Index), 적응 능력 지수(ACI, Adaptive Capacity Index)로 구성하였다. 침수취약성을 구성하는 대리변수는 선행연구에서 침수피해액 또는 인명피해에 유의한 영향 관계가 있는 변수들을 중심으로 선정하였다. 재해노출 지수는 기후적 특성을 갖는 변수들로 구성하고, 민감도 지수와 적응능력 지수는 변수의 특성에 따라 물리적 요인, 경제적 요인, 인구적 요인으로 구분한다. 이때 변수들은 2019년을 시간적 기준으로 하되, 재해노출과 관련된 대리변수는 2015년부터 2019년까지 5년간 평균값으로 한다.

침수취약성 요인 중 재해노출 지수는 기상적 요인 특성을 갖는 '일 최대 강수량'과 '일 강수량 80mm 이상인 집중호우 일수'로 구성한다. 강수량이 많고, 단기간에 많은 양이 내리는 집중호우가 발생할수록 침수 발생 노출 영향이 커짐에 따라 선행연구에서 대표적으로 사용되는 대리변수이다.

민감도 요인 중 물리적 특성을 나타내는 변수로는 '노후 단독주택 비율'과 '1인당 불투수 면적', '하천면적 비율', '급경사지 밀도'를 사용하였다. '노후 단독주택 비율'은 전체 주택 중에 준공 30년 이상 노후 단독주택 비율로, 「도시 기후변화 재해 취약성 분석 지

침」상에서 제시하는 자연재해 취약주택에 해당하므로 물리적 민감요인으로 반영한다. '시가화 면적 비율'은 전체 행정면적 중 도시지역 면적 비율로 건물, 콘크리트, 아스팔트 등으로 덮여 빗물이 침투하지 못하는 불투수 면적을 나타내며, 침수 발생을 가중시키는 요인으로 알려져 있다. '하천면적'은 시·군·구별 행정면적 대비 하천면적 비율을 나타내고, '급경사지 밀도'는 국민재난안전포털에 급경사지로 공개된 시·군·구별 급경사지의 행정면적당 밀도이다. 이들은 폭우 발생 시, 하천범람에 따른 침수 발생 가능성과 급경사지의 빠른 우수유출 속도로 붕괴 위험이 높기 때문에 물리적 민감요인으로 작용한다. 경제적 특성으로는 '기초생활수급자 비율'을 고려하는데, 기초생활수급자는 재산, 소득이 적은 경제적 취약계층으로, 경제적 취약성을 대표할 수 있다. 한편, 인구적 특성으로 '영유아 및 노령자 비율'을 고려하는데, '영유아 및 노령자 비율'은 전체 인구 중에서 5세 이하와 65세 이상 인구 비율로 신체적 취약성을 나타낸다.

적응역량 요인 중 물리적 특성 변수로는 '계획인구 대비 하수관거 용량'과 '시가화 면적 대비 녹지면적', '1인당 도시공원면적', '우수지 및 저류시설 면적 비율', '제방 면적 비율'을 고려한다. '하수관거 시설용량'은 계획인구 대비 하수관거 시설용량으로 불투수 면적의 빗물 배수를 빠르게 처리할 수 있는 역량을 반영한다. '시가화 면적 대비 녹지면적'은 「국토의 계획 및 이용에 관한 법률」에서 지정하는 도시지역 면적 대비 녹지지역 면적 비율로, 도시지역의 빗물 투수율을 높일 수 있는 요인으로 작용한다. 한편, 도시공원 확보기준으로 활용되는 '1인당 도시공원 면적'은 도시지역 내 생활권 공원과 주제공원의 면적을 인구수로 나눈 값으로, 인구 거주지 내 빗물 투수율을 높이는 요인으로 반영한다. '우수지 및 저류시설 면적 비율'과 '제방 면적 비율'은 행정면적 대비 도시계획시설 면적 비율로, 하천의 범람을 방지하고 빗물 유출량을 저감시키는 요인으로 작용한다. 경제적 특성으로 '지방재정 자립도'는 방재시설 사업 및 안전교육 등 안전환경 조성을 위한 지자체의 자체적인 재정 투입 역량을 반영하였다. 인구적 특성을 갖는 '인구 대비 의료인력 수'는 침수로 인한 인명피해 발생 시 응급조치 및 의료행위로 사망 및 부상자를 치료할 수 있는 인적 대응력을 나타낸다. '119안전센터 밀도'는 행정면적 내 119안전센터 수로 사고 발생지역까지 119안전센터의 접근성을 나타내며, 높은 접근성은 사고 발생 시 소방구급대원들의 신속한 구급 대응력을 반영한다.

IV. 연구 결과

1. 침수피해와 침수취약요인 기초통계 분석

1) 우리나라 침수피해 특성

피해액과 인명피해 간 산점도 분포(그림 2)와 기초통계분석(표

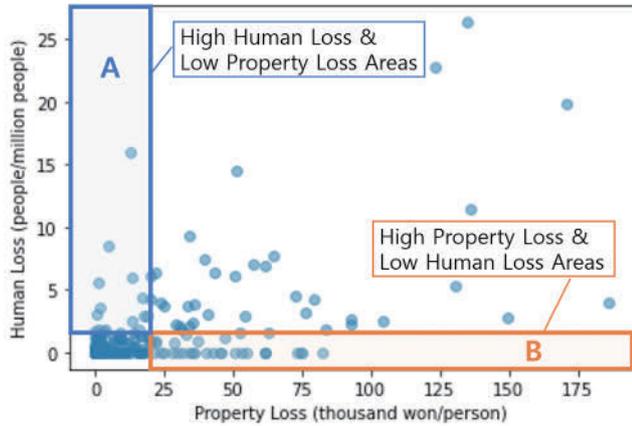


Figure 2. Scatter plot between property damage and human casualties

Table 3. Basic statistical analysis for flood damage

Category	Obs.	Mean	Std. dev.	Min.	Max.
Avg. property damage (1,000 won/person)	229	20.2664	31.7924	0	186.0118
Avg. human casualties (people/1,000,000 people)	229	15.4723	33.9666	0	263.6168

3)을 살펴보면, 피해액 평균은 20,2664(천 원/인)이고, 표준편차는 31,7924(천 원/인)인데, 인명피해 평균은 15,4723(인/백만인)이고, 표준편차는 33,9666(인/백만인)으로 인명피해 데이터의 평균으로부터 분산이 피해액 데이터보다 큼을 알 수 있다.

침수피해 지표의 2010년부터 2019년까지 10년간 229개 시·군·구별 평균값은 다소 편향된 자료 분포를 나타낸다. 대다수의 자료값은 0값에 가깝고 평균값이 낮게 나타나는 반면, 일부 지역에서는 매우 극단적인 피해액과 인명피해가 나타난다. 예를 들어, <그림 2>에서 인명피해는 없지만 피해액이 발생했거나(B section), 피해액은 거의 발생하지 않았지만 인명피해는 높게 나타난 지역(A section)이 이에 해당한다. 이와 같이, 피해액과 인명피해의 발생 양상은 반드시 일치하지 않음을 알 수 있다.

본 연구에서는 피해액과 인명피해 간의 상관관계를 알아보기 위해 Kendall's tau의 순위 상관관계 분석을 실시하였다. Kendall's tau의 순위 상관계수(τ)는 두 변수 간 관계가 비선형적인 관계일 때 활용되는 방법으로, 한 변수가 증가할 때 다른 변수도 증가하는 횟수와 감소하는 횟수 전체를 측정하여, 그 차이로 상관관계를 분석하는 것이다. 상관계수는 -1에서 1까지 범위를 가질 수 있고, 양의 상관관계의 경우 0.4에서 0.7까지는 상관관계가 있다고 보며, 0.2에서 0.4까지는 약한 상관관계가 있다고 본다. 분석 결과, 피해액과 인명피해 지표 간 Kendall's tau 순위 상관계수는 0.2489 정도로 약한 양의 상관관계를 갖는 것으로 나타났다. 이는 두 지표 간에 관계가 있으나 관계의 강도가 매우 높지는 않

음을 알 수 있다. 이와 같은 결과는 기존 다수의 연구에서 피해액만을 피해지표로 고려하고 있는 것이 침수피해 현황을 분석하고 대응책을 마련하는 데에 충분하지 않음을 시사한다.

2) 침수취약요인의 기초통계 분석

본 연구에서 설명변수로 활용되는 침수취약성 지표의 단위와 기초 통계량은 <표 4>와 같다. 재해 노출 요인 '일 최대 강수량'은 연평균 118.18mm의 값을 나타내 집중호우 기준인 일 최대 강수량 80mm를 넘는 것을 알 수 있다. '집중호우 일수'의 경우 최솟값은 0.4일이고 최댓값은 6.0일로 집중호우 발생 일수가 낮은 지역과 높은 지역의 편차가 크다.

민감도 요인 중 물리적 특성에 해당하는 '노후 단독주택 비율'은 최솟값이 0.67%인 데 비해, 최댓값은 69.02%로 나타나는데, 이는 우리나라 주거특성과 관련이 있다. 통계청³⁾에 따르면 우리나라 주택 유형 중 아파트(60.6%)와 단독주택(23.1%)의 점유율이 높는데, 아파트의 경우 수도권과 광역 대도시권 주택 유형의 대다수를 차지하고, 단독주택의 경우 전라남도과 경상남·북도 지역에서 점유율이 높게 나타난다. '시가화 면적 비율'은 서울을 비롯한 6대 광역시와 경기도는 83.64%(%)로 전국 평균값보다 높은 값을 보이지만, 강원도를 비롯한 6개 각 도와 특별 자치시·도의 평균값은 13.79%(%)로 매우 낮은 값을 보인다. 이는 '시가화 면적 비율'이 도시지역에서 더 크게 나타남에 따라 도시지역의 침수 민감도가 높음을 나타낸다.

한편, 적응역량 요인 중 '인구 대비 하수관거 시설용량'은 평균값이 3.77(천 톤/일·인)인데, 최솟값은 0.02(천 톤/일·인), 최댓값은 93.48(천 톤/일·인)를 보인다. 서울시와 경기도의 하수관거 시설용량 평균값은 각각 5.08(천 톤/일·인), 8.73(천 톤/일·인)를 보여 매우 높은 값을 보이는 반면, 강원도를 비롯한 6개 도와 특별자치시·도의 평균값은 2.79(천 톤/일·인)에 그친다. 이는 인구 대비 빗물을 포함한 하수 처리 역량이 수도권 지역에서 상대적으로 높음을 보여준다. 반면, 도시지역의 투수 역량과 관련이 있는 '시가화 면적 대비 녹지면적'은 64.75%(%)의 평균값을 보이고, '인당 도시공원 면적'은 19.5(m²/인)의 평균값을 보인다. 녹지지역과 도시공원 두 지표에 대해 전국 평균값보다 높은 평균값을 보이는 지역은 강원도와 전라남·북도, 경상남·북도로 나타났다. 반면, 서울시와 대구시, 인천시는 녹지지역과 도시공원 두 지표에 대해 전국 평균값보다 낮은 평균값을 보이며, 투수면적이 낮아 침수피해 위험이 높음을 알 수 있다.

2. 침수피해에 따른 주요 침수취약요인 분석

1) 기계학습 모형

두 번째 연구 질문인 '피해액과 인명피해 발생에 기여하는 주요한 침수취약요인의 차이는 무엇인가?'에 대한 답을 찾기 위해 본

Table 4. Basic statistics and feature importance ranking

Category	Indicators	Variable name	Unit	Obs.	Basic statistics				Feature importance		
					Mean	Std. Dev.	Min.	Max.	Property loss	Human casualties	
EI	Climate factor	Average of maximum daily precipitation	El_rainfall	mm	229	118.18	26.39	71.5	206.1	○	○
		Average of days over 80mm rainfall	El_rainday	day	229	1.85	0.80	0.4	6.0	◎	○
SI	Physical factors	Share of old single houses	SI_oldsingle house	%	229	19.87	15.25	0.67	69.02	◎	◎
		Share of urbanized area per administrative area	SI_urbanarea_%	%	229	46.27	42.83	0.33	100.0	◎	○
		Share of river area	SI_riverarea	%	229	2.08	3.25	0.0	23.93		
	Economic factors	Density of steep slope	SI_slope	points/km ²	229	0.12	0.14	0.0	1.06	○	○
		Share of basic livelihood recipients	SI_poor	%	229	4.37	1.54	1.22	9.58	○	◎
		Share of old and young people	SI_oldyoung	%	229	23.26	6.87	8.05	40.21	○	◎
ACI	Physical factors	Daily sewage pipe capacity per person	ACL_capacity sewage_pop	thousand ton/day-person	229	3.77	7.38	0.02	93.48	○	◎
		Share of green area in urban area	ACL_green_%	%	229	64.75	21.24	0.0	92.84		○
		Urban park area per person	ACL_urbanpark	m ² /person	229	19.50	18.42	0.0	130.60		
		Share of detention basin and water retention facility area	ACL_reservoir	%	229	0.11	0.31	0.0	3.30		
	Economic factors	Share of embankment area	ACL_embarkment	%	229	0.27	0.29	0.0	2.11	○	○
		Local financial independence	ACL_finance	%	229	20.38	12.83	4.0	68.9	◎	
		Population factors	Share of medical personnel per thousand people	ACL_doctor	people/thousand people	229	1.90	2.22	0.44	24.56	○
Density of 119 safety center	ACL_119		points/km ²	229	0.08	0.13	0.0	1.06	○	○	

◎ : Top 4 feature importance rankings
 ○ : Between 5th and 8th feature importance rankings
 ○ : Between 9th and 12th feature importance rankings

연구에서는 기계학습 모형을 활용한다. 이는 기계학습 모형의 특성으로 이상적인 사건이 발생하는 자연재해 분야에서 기계학습 모형의 정확도가 통계 모형보다 더 높은 것으로 알려져 있기 때문이다(Tehrany et al, 2014; 이상혁 외, 2016; 최창현 외, 2018; Chen et al., 2020). 이에 본 연구에서는 기존 전통적인 회귀모형으로 침수피해에 영향을 미치는 침수취약요인을 분석한 연구에서 나아가, 기계학습 모형을 활용하여 침수피해와 취약요인 간 영향 관계를 살펴보고자 한다. 이를 위해 선형회귀 모형과 대표적인 기계학습 모형 4가지의 모형에 대해 평가하여 좋은 성능을

보이는 모형을 채택한다. 기계학습 모형은 전체 데이터를 학습 표본과 평가 표본으로 분할하여, 학습 데이터를 통해 모형 추정 후, 평가 데이터의 실제값과 모형에 의해 예측된 값 간의 차이를 통해 모형의 설명력과 정확도를 평가한다. 본 연구에서는 회귀모형을 활용하여 피해액과 인명피해를 추정하기 때문에, 회귀모형의 평가 지표로 실제값의 분산 대비 예측값의 분산 비율을 나타내는 R²(식 (1))과 예측값과 실제값 오차의 합을 보는 RMSE(식 (2))와 MAE(식 (3))을 활용한다. 따라서 R²는 높고, RMSE와 MAE는 작은 모형을 성능이 좋은 모형으로 채택한다.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{3}$$

본 연구에서는 학습 표본은 전체 데이터의 80%인 183개, 평가 표본은 20%인 46개로 할당하고, 모형학습 시 과적합을 방지하기 위해 각각의 모형에 필요한 매개변수를 조정하였다. 모형별 설명력과 정확도 평가 결과는 다음 <표 5>와 같다. 모형 평가 결과, 피해액과 인명피해에 대한 기계학습 모형은 선형회귀 모형보다 높은 설명력과 정확도를 보인다. 따라서 선형적인 관계를 가정하는 선형회귀 모형보다 비선형적인 관계를 포착할 수 있는 기계학습 모형이 더 적합함을 알 수 있다.

피해액 모형의 경우 Gradient Boosting Regression이 R^2 는 64.05%로 가장 높고, RMSE와 MAE는 각각 16.91, 10.17로 가장 낮다. 반면 인명피해 모형 평가에서는 Random Forest Regression 모형이 R^2 가 31.49%로 가장 높고, RMSE 27.82와 MAE 13.54로 가장 낮아 해당 모형을 사용하였다. 피해액과 인명피해의 R^2 값에 차이가 크게 나타나는 이유는 자료 분포의 특성 차이 때문이다. 피해액에 비해 인명피해의 자료 분포는 0 값이 많은 특성으로 일반적인 선형회귀 모형으로 추정이 어렵다. 트리 기반 기계학습 모형은 데이터 집단을 분할하는 규칙에 의해 반복 학습하고, 이때 집단을 분할하는 기준은 예측치와 실측치의 잔차를 최소화하는 특성 값이며, Regression 모형은 분할된 집단 내에서 선형회귀 모형으로 예측값을 추정하는 방식이기 때문에 인명피해 추정 모형의 설명력은 피해액 추정 모형에 비해 상대적으로 낮을 수밖에 없다.

Random Forest와 Gradient Boosting은 모두 트리 기반 모형으로 훈련 데이터에서 여러 개의 샘플을 생성하고, 각 샘플에서 의사결정 나무를 통해 예측된 결괏값 중 가장 많은 값으로 예

측하는 방식이다. 무작위로 추출된 여러 개의 훈련 데이터 내에서 임의의 선택된 특성으로 예측된 값을 집계하기 때문에, 모형이 훈련 데이터에 과도하게 맞춰 추정되는 과적합(overfitting)을 방지할 수 있고, 예측력이 높다. Gradient Boosting model은 Random Forest와 달리 이전 트리의 오차를 보완하는 방식으로 트리를 만드는 차이점이 있다(Ha and Kang, 2022; 박준상·이수기, 2022).

본 연구에서는 Python 내 대표적인 기계학습 라이브러리인 Scikit-learn에서 제공하는 Gradient Boosting Regressor와 Random Forest Regressor 앙상블 학습방법을 활용하여 모형을 구축하였다. Gradient Boosting model의 주요 hyper parameter로는 부스팅 반복수를 결정하는 n_estimators, 이전 트리의 오차를 얼마나 강하게 보정할지 제어하는 learning_rate, 각 학습 샘플에 분할되는 표본의 비율을 나타내는 subsample, 트리의 최대 깊이를 제한하는 max_depth 등이 있다. 한편, Random Forest Regression 모형의 주요 hyper parameter로는 결정 트리의 개수를 지정하는 n_estimators, 노드를 분할하기 위한 최소 샘플 데이터 수를 지정하는 min_sample_split, 최적의 분할을 위해 고려할 최대 특성 개수인 max_features, 트리의 최대 깊이는 정하는 max_depth 등이 있다. 본 연구에서 최적화된 hyper parameter의 값을 찾기 위해 GridSearchCV 모듈을 활용하였다. GridSearchCV는 hyper parameter 후보군 중에서 가장 높은 성능을 갖는 조합을 찾고, 이를 훈련 데이터 내에서 구간을 분할하여 교차 검증하는 방법이다. 이를 통해 도출된 최적의 hyper parameter 값으로 피해액 모형은 n_estimators = 100, learning_rate = 0.04, subsample = 0.5, max_depth = 10이고, 인명피해 모형은 n_estimators = 300, min_samples_split = 10, max_features = 2, max_depth = 10으로 설정하였다.

2) 침수피해에 영향을 미치는 주요 침수취약요인

Gradient Boosting과 Random Forest 모형에서 침수 취약요인과 예측된 침수피해 간 관계를 해석하기 위해 Python 내 해석 패키지인 SHAP(SHapley Additive exPlanations)을 이용하

Table 5. Machine learning model accuracy

Model	Property damage			Human casualties		
	R^2	RMSE	MAE	R^2	RMSE	MAE
Linear Regression	0.4965	22.51	14.8359	0.2172	29.9869	16.7414
Random Forest	0.6274	22.2272	10.9758	0.3149	27.8187	13.5386
Gradient Boost	0.6405	16.9111	10.174	0.2748	30.171	18.9952
Cat Boost	0.5488	23.9248	12.2997	0.2931	28.851	17.0757
XGBoost	0.5604	23.2672	14.3658	0.2574	32.9621	17.7255

*p-value<0.01

였다. SHAP에서 각 특성의 기여도 산출을 위해 사용하는 shapley values란 게임이론을 바탕으로 게임에서 각 플레이어의 기여도를 계산하는 방식이다. 즉, 특성별 기여도를 산출하기 위해 여러 특성 조합을 구성하고, 해당 특성의 유무에 따라 예측값의 변화량을 평균하여 나타낸 것이다. 본 연구에서는 트리 앙상블 모델에 적합한 Tree SHAP을 활용하여 Gradient Boosting Regression과 Random Forest Regression 모델을 해석하였다.

Shapley value 값을 이용하여 모형에 활용된 특성의 중요도(feature importance)를 산출한다. 해당 특성에 대해 예측치별 shapley value의 절댓값 크기로 산정하는 방식을 Local feature importance라 하고, 이를 각 특성에 대해 절댓값의 평균으로 크기를 나타낸 것이 Global feature importance라 한다. 본 연구에서는 특성 변수별 값의 크기에 따라 예측값에 미치는 영향의 일반적인 경향성을 해석하기 위해서 Local feature importance를 해석하지 않고, 이를 종합한 Global feature importance를 해석한다. 본 연구에서는 <표 6>과 같이, Global feature importance에서 특성 값과 예측치에 기여도(Shapley value) 간 양(+)의 상관관계가 있는 경우 붉은색으로 표현하고, 음(-)의 상관관계가 있는 경우 푸른색으로 표현하였다. 일반적으로 Global feature importance는 해석의 유의미성을 고려하여 상위 3-4개 변수의 중요도와 영향의 방향성에 대해 해석한다(박준상·이수기, 2022; 이제승·이현희, 2019).

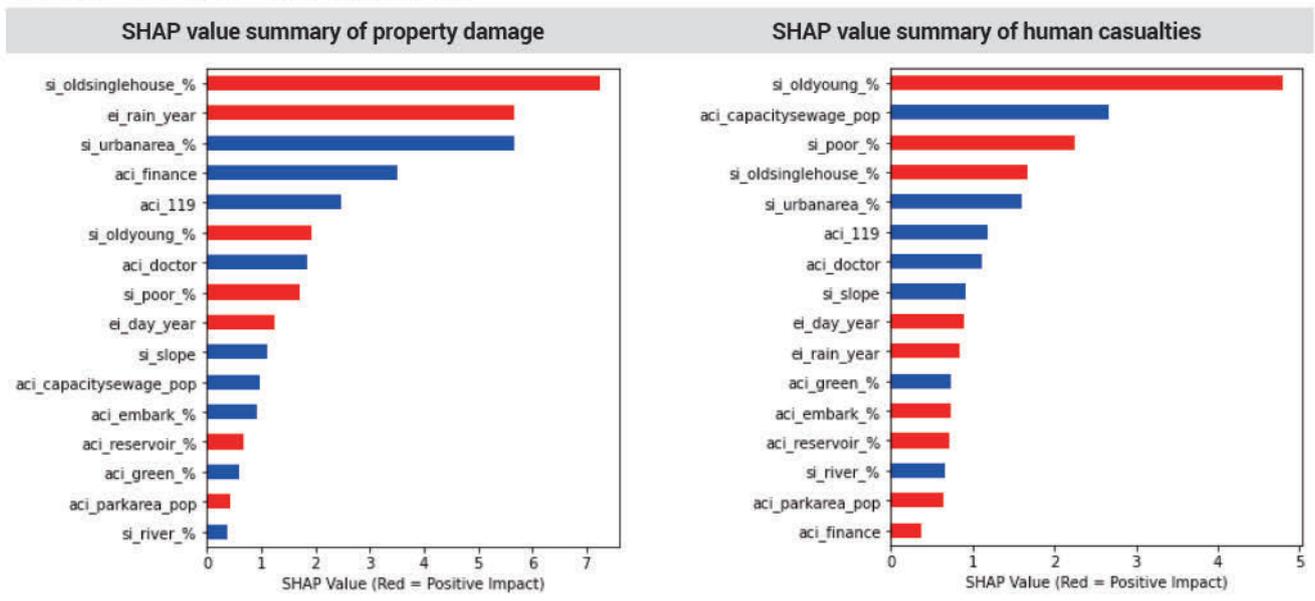
피해액에 영향을 미치는 중요한 특성으로 '노후 단독주택 비율(si_oldsingehouse_%)'과 '일 최대 강수량(ei_rain_year)'은 양(+)의 방향으로 영향력이 매우 높게 나타났다. 준공 30년 이상 노후 단독주택은 내수배제 시설의 노후화로 강우 노출시 내수침수가 발생할 가능성이 높기 때문이다. 이와 더불어 강한 강우 노출은 내·외수 침수로 인한 시설피해를 유발하고, 그로 인한 피해액

이 크게 발생하게 하는 영향이 있다.

반면, '시가화 면적 비율(si_urbanarea_%)'과 '재정 자립도(aci_finance)'는 피해액 발생에 음(-)의 상관관계를 가지며 중요한 영향요인으로 나타난다. 지난 10년간의 재해 피해액 발생의 76.5%가 공공시설 피해이며(행정안전부, 2020), 지난 5년간 발생한 침수발생 상세원인으로 태풍 및 집중호우 발생, 배수용량 부족, 하천 범람 및 해수면 상승 등이 두드러진다.⁴⁾ 이를 고려하면 공공시설 피해액은 많은 부분 하천 및 수리, 사방시설과 관련하여 발생했을 것으로 추정되고, 이와 같은 해당 도시계획시설은 도시지역보다 비도시지역(농림지역, 관리지역, 자연환경보전지역 등)의 용도지역으로 지정되어 관리되는 특성이 있다. 따라서 용도지역상의 비도시지역 비율이 높을수록, 하천 및 수리, 사방 시설에서 발생하는 공공시설 침수피해액이 크게 발생하는 영향이 있는 것으로 해석된다. 재정 자립도의 경우, 지자체의 재난안전예산은 국고보조금에 대한 의존도가 높고, 국고보조금을 매칭하기 위해 순수 지자체 사업예산이 필요하기 때문에 침수피해와 관련된다(한국행정연구원, 2020). 즉, 재정 자립도가 높은 지자체는 재난안전예산 확보가 용이하기 때문에 재난취약지역·시설의 위험 요소 제거 및 기능 회복, 재난안전 관리 목적의 시설 구축 및 기능 강화를 위한 예산 투자가 상대적으로 높다. 그 결과 재정 자립도가 낮은 지자체는 침수피해가 높은 영향이 있다.

한편, 인명피해 증가에 영향을 미치는 중요한 특성으로 '영유아 및 노령자 비율(si_oldyoung_%)'과 '인구당 하수관거 시설용량(aci_capacitysewage_pop)', '기초생활수급자 비율(si_poor_%)', '노후 단독주택 비율(si_oldsinglehouse_%)' 요인이 높게 나타났다. 영유아 및 노령자 비율과 기초생활수급자 비율은 인명피해와 양(+)의 상관관계가 있는 것으로 나타난다. 신체적 취약인구는 사고 발생 시 상황 판단 및 재해 대응 역량이 부족하

Table 6. Result of SHAP value summary plot



고, 경제적 취약인구는 저렴한 지가로 침수취약 지역에 거주할 가능성이 높기 때문에 본다(Tellman et al., 2020; Alderman et al., 2012; Jonkman and Kelman, 2005; Lowe et al., 2013). 또한 노후 단독주택 비율이 높은 지역에서 인명피해가 크게 발생하는 영향이 있는데, 이는 노후 단독주택이 낮은 주거비용으로 인해 취약계층의 거주 비율이 높게 나타나고, 침수에 민감한 주택 유형에 해당하기 때문이다(서정석·한우석, 2019).

반면, 하수관거 시설용량은 인명피해와 음(-)의 상관관계가 나타나므로써 하수관거 시설용량이 증가할수록 인명피해가 감소하는 영향이 있음을 의미한다. 침수발생 상세원인으로 배수용량 부족이 빈번하게 언급되듯이, 강한 강우로 우수처리 용량 급증 시, 배수 처리용량의 한계는 돌발홍수를 유발하기 쉽다. 그리고 돌발홍수는 침수로 인한 인명피해의 주요 원인으로 작용하기 때문에(Jonkman, 2014) 하수관거 시설용량은 인명피해 발생에 주요한 영향을 미친다.

3. PDP를 활용한 비선형관계 분석

세 번째 연구 질문인 ‘피해액과 인명피해에 대한 침수취약성 요인의 비선형적 영향력은 어떻게 다른가?’에 대한 답을 찾기 위해 본 연구에서는 기계학습 모형을 해석하기 위한 다른 방법인 PDP(Partial Dependence Plot)를 활용하였다(〈표 7〉 참조). 이는 입력 특성이 예측값에 미치는 영향을 시각적으로 분석하기 위한 방법으로, 예측값에 대한 입력 특성의 한계 효과(marginal effect)를 그래프로 확인할 수 있다. PDP는 각 변수 간 상관관계가 없음을 가정하고, 관심 입력 특성을 제외한 나머지 입력 특성은 평균값으로 고정했을 때, 관심 입력특성의 값에 따라 변화하는 예측치의 분포를 나타낸다(Ha and Kang, 2022). 이때, 특정 입력 특성에서 표현되는 예측치는 특정 입력특성 값을 모든 샘플에 동일하게 적용하였을 때 예측되는 값의 평균값이기 때문에, 샘플 간 성질이 매우 다를 경우, 예측값이 상쇄되어 표현되는 한계가 있다. 본 연구에서는 Python 내 해석 패키지인 Dalex를 이용하여 PDP를 그리고, 입력 특성과 예측변수 간 선형관계 및 비선형관계를 분석하였다.

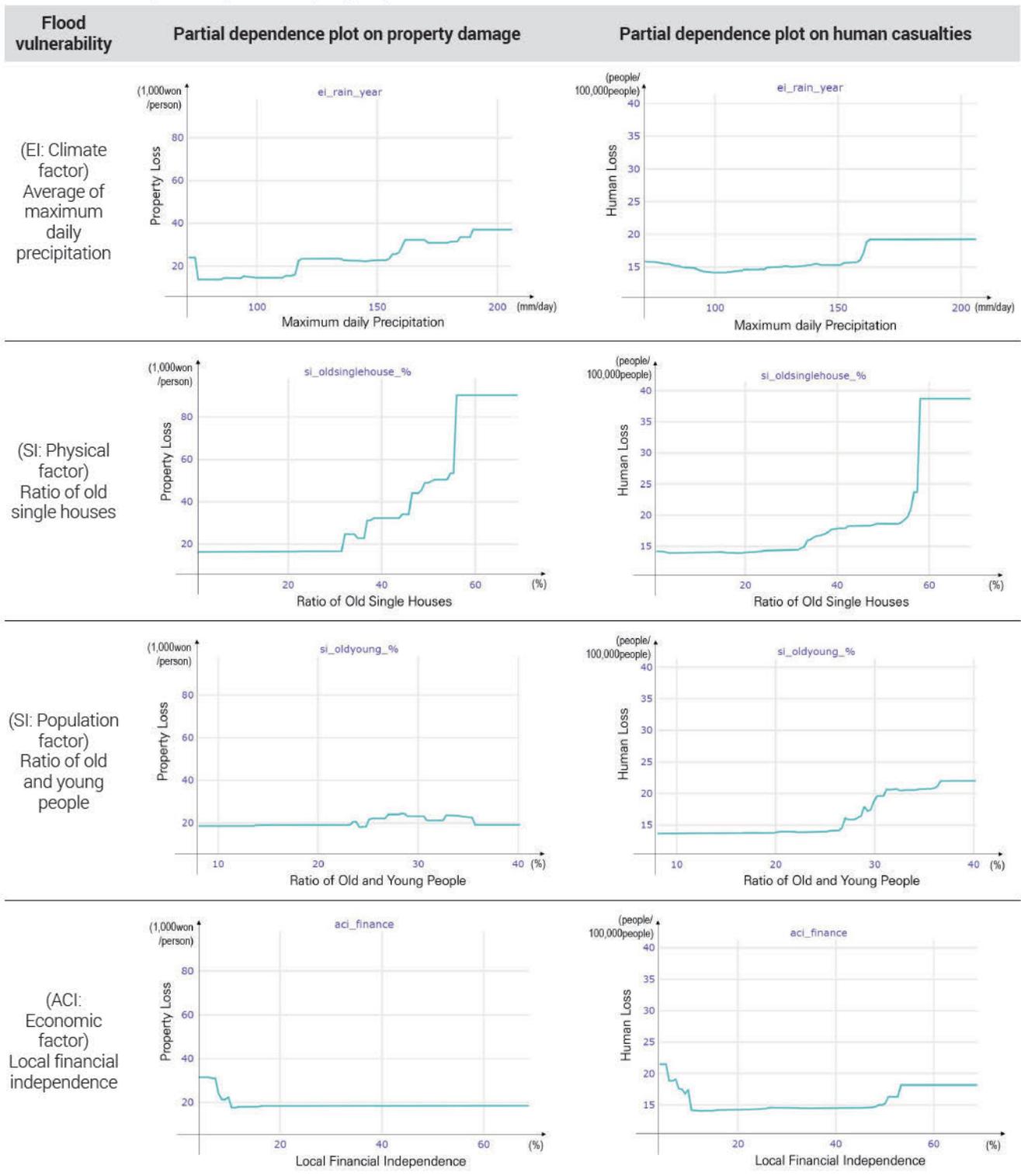
본 연구에서는 침수피해에 대하여 변수 중요도가 높고, 비선형적인 관계를 나타내는 침수취약성 요인(EI, SI, ACI)별 대표적인 침수취약 변수에 대해 PDP를 해석하고자 한다. 앞의 Feature importance plot에서는 변수 중요도를 중심으로 피해액과 인명피해의 중요변수의 차이를 해석하였다면, PDP에서는 같은 지표이지만 침수피해액과 인명피해에 대한 변수의 부분 의존도에 차이를 비교하기 위해 침수취약성 요인별로 대표적인 변수를 선별하였다. 피해액에 대한 부분 의존도(Partial Dependence)는 20(천 원/인)부터 100(천 원/인)의 범위를 갖는 반면, 인명피해에 대한 부분 의존도(Partial Dependence)는 15(인/십만인)에서

40(인/십만인)의 범위를 갖는다. 두드러지는 PDP 특성을 갖는 침수취약요인을 중심으로 해석하면 다음과 같다. 먼저, ‘일 최대 강수량’은 피해액에 대해 120(mm/일)과 190(mm/일) 사이에서 단계적으로 영향력이 높아지는 것을 알 수 있다. 반면, 인명피해에 대해서는 160(mm/일) 이상에서만 영향력이 높게 나타난다. 이는 연평균 일 최대 강수량 160(mm/일)에서 피해액과 인명피해 모두 높게 발생함을 의미하고, 이는 서울에 기록적인 폭우가 내렸던 2022년 8월 8일에서 9일에 서울 한강 이남 지역에 발생한 집중강우가 24시간 지속 최대 강우량 435(mm/일)로 관측된 점을 고려하면,⁹⁾ 그 당시 강우량의 36.78%에 해당하는 강우 발생에서도 피해액과 인명피해가 높게 발생하는 영향이 있음을 알 수 있다. 한편, 2020년 8월 7일과 8일 이틀 동안 영산강과 섬진강 유역 일대에 발생한 홍수의 경우, 장성군에서 24시간 지속 최고 강우량 316.6(mm/일), 남원시에서 365.0(mm/일)로 나타났다(이경훈 외, 2020). 이를 통해 비도시지역은 이 당시 강우량의 약 46.95%에 해당하는 강우 발생에서도 심각한 침수피해가 발생할 수 있음을 알 수 있다.

‘노후 단독주택 비율’의 경우 피해액에 대해서 56% 이상의 지역에서 매우 극명하게 피해액이 크게 발생하고, 인명피해에 대해서도 58% 이상의 지역에서 매우 높은 영향력이 나타난다. 2019년을 기준으로 전국 평균 노후 단독주택의 비율은 59.61%로 나타나는데, 7개 시도의 평균은 70% 이상의 비율을 보이는 반면, 서울시와 경기도, 광역시는 전국 평균값보다 낮은 노후 단독주택 비율을 보인다. 이를 통해 다른 영향요인을 통제했을 때, 노후 단독주택 비율이 60% 이상으로 높은 지역은 피해액과 인명피해가 모두 높게 발생하는 영향이 있고, 이러한 지역은 대도시 지역을 제외한 지역임을 알 수 있다. 또한 노후 단독주택 비율이 30%와 60% 전후에서 단계적으로 피해액과 인명피해에 대한 부분 의존도가 높아지는 비선형적 관계를 보이므로, 지자체별 노후 단독주택 비율에 따라 단계별 침수피해 예방 정책을 제시할 필요가 있음을 시사한다.

민감 요인의 인구적 특성인 ‘영유아 및 노령자 비율’은 피해액보다 인명피해에 부분 의존도가 더 분명하게 나타난다. 피해액은 영유아 및 노령자 비율 25%에서 35% 사이의 지역에서 피해액 발생이 약간 증가하는 영향이 있는 반면, 인명피해는 영유아 및 노령자 비율 32% 이상인 지역에서 인명피해 발생이 1.3배 높게 나타난다. 이를 통해 인명피해가 피해액보다 영유아 및 노령자 비율의 영향을 더 많이 받는다는 것을 알 수 있다. 또한 영유아 및 노령자 비율이 25% 이상에서 인명피해에 대한 부분 의존도가 급격하게 높아지는 비선형적 관계를 보이므로써, 인명피해 취약 지표로 영유아 및 노령자 비율을 고려 시 영향 임계점에 대한 기준을 제시할 필요가 있음을 시사한다. 한편, 피해액에 대한 영유아 및 노령자 비율의 부분 의존도가 35% 이상의 지점에서 다시 감소하는 것은 경제활동인구가 65% 이하로 낮은 지역의 물리적 자산

Table 7. Result of partial dependence plot (PDP)



가치가 상대적으로 낮아 피해액이 낮게 산정된 결과로 해석된다. 마지막으로 경제적 적응역량 요인인 '재정 자립도'는 피해액과 인명피해에 대해 각각 약 10%와 약 15% 이하인 지역에서 공통적으로 약 1.5배 높은 영향력을 보인다. 이는 재정 자립도 특정 기준치(약 10%) 이하의 지역에서 피해액과 인명피해 모두 높게 발생하는 특성이 있음을 보여준다. 따라서 특정 기준값 이하로 낮은

재정 자립도를 보이는 지자체에 대해 재난안전예산 관련 국고보조금 지원 기준을 조정할 필요가 있다. 국고보조금 대비 매칭되는 지자체 예산 비율을 낮게 조정하면, 재정 취약 지자체에 대해 국고보조금 지원을 확대할 수 있고, 이를 통해 지자체에서 재해 취약지구에 대한 사전 보수 및 방재기능 강화, 재해예방사업을 등을 시행할 수 있을 것이다.

V. 결론 및 시사점

기존 국내 선행연구에서는 침수로 인해 나타나는 피해액과 인명피해가 나타나는 양상이 다르고, 각 침수피해에 영향을 미치는 요인들이 다름에도 불구하고 피해액을 중심으로 침수 관련 연구가 이루어졌다. 더욱이, 침수피해는 일상적으로 관측되는 값이 아닌, 특정 시점과 특정 지점에서 극단적으로 큰 값으로 나타나는 특징이 있음에도 불구하고 전통적 회귀모형이 주로 사용되어 왔다. 침수피해 자료의 특성을 고려할 때, 비모수적, 비선형적 추정방식인 기계학습 모형이 침수피해와 침수취약요인 간 관계를 설명하기에 적합한 분석방법이 될 수 있다. 이와 같은 배경에서 본 연구는 침수피해를 나타내는 지표로 피해액뿐만 아니라 인명피해자 수를 고려하여 침수피해 지표의 특성을 분석하고, 기계학습 모형을 통한 침수피해 결정요인을 분석하여 기존 연구에서 부족한 점을 보완하고자 하였다. 주요한 분석 결과는 다음과 같다.

첫째, 국내에서 발생한 피해액과 인명피해자 수 간 발생 양상은 약한 상관관계를 가지고 있는 것으로 나타났다. 이는 동일한 재해 노출 상황에서도 발생하는 침수피해 양상이 동일하지 않음을 시사한다. 기존 연구에서 피해액을 주요한 지표로 침수피해를 산정하는 점을 고려할 때, 피해액은 낮지만 인명피해가 높은 지역에 대한 침수피해 인식이 부족할 수 있었음을 의미한다.

둘째, 모형 검증을 통해 선형회귀 모형과 앙상블 기계학습 모형을 비교한 결과, 선형회귀 모형보다 앙상블 기계학습 모형의 설명력과 정확도가 전반적으로 높게 나타났다. 이는 기존 선형회귀 모형으로 침수피해 영향요인을 밝혔던 연구보다 기계학습 모형을 통해 침수피해와 침수취약요인 간 관계를 밝히는 것이 비선형적 관계까지 설명할 수 있어 더 적절한 모형임을 의미한다.

셋째, 피해액과 인명피해에 대해 영향을 미치는 침수취약요인은 상이함을 밝혔다. 피해액을 증가시키는 주요 침수취약요인은 높은 노후 단독주택 비율과 재해 노출, 높은 비도시지역 비율 및 낮은 재정 자립도 요인으로 나타났다. 반면, 인명피해를 발생시키는 주요 침수취약요인은 높은 영유아 및 노령자 비율과 낮은 하수관거 시설용량, 높은 경제적 취약인구 비율, 높은 노후 단독주택 비율로 나타났다. 두드러지는 차이점으로 피해액은 물리적 취약시설과 재해 노출의 영향을 크게 받는 반면, 인명피해는 신체적·경제적 취약인구 거주율과 물리적 시설 대응력의 영향을 크게 받는다는 것이다. 이는 피해액과 인명피해 대응을 위해 침수피해 지표별 서로 다른 기준으로 취약지역을 진단할 필요가 있음을 시사한다.

이와 같은 연구 결과는 노후주택과 지역의 경제력이 침수피해액을 가중시킨다(Jeong and Yoon, 2018; Yoon, 2012; 서만훈 외, 2016)는 결과와 유사한 결과이다. 반면, 본 연구에서는 기존 연구와는 달리 비도시지역 비율이 높을수록 피해액이 크게 발생하는 것으로 나타났는데, 이는 주된 침수피해액 발생 원인인 하

천 및 수리시설이 비도시지역에 주로 위치하기 때문인 것으로 해석될 수 있다. 한편, 우리나라의 경우 인명피해에 영향을 미치는 요인에 대한 실증연구가 많지 않지만, 해외 연구에서 나타난 주요 영향요인과 유사한 결과가 도출되었다. 해외 실증연구에서도 신체적·경제적 취약인구가 침수피해에 취약하고, 물리적 대응역량이 인명피해를 저감하는 주요 특성으로 나타났다(Doocy et al., 2013; Jonkman, 2014; Zahran et al., 2008).

마지막으로, 침수피해와 침수취약요인 간 비선형적 관계 분석을 통해 침수 취약요인의 피해에 대한 영향은 단계적으로 나타나거나 특정 임계점이 존재하는 것으로 나타났다. 또한 동일한 침수취약요인에 대해서도 피해액과 인명피해에 따라 다른 임계점을 보였으며, 부분 의존도의 정도에 차이가 존재하였다. 따라서 향후 연구에서는 침수취약지역의 평가를 위해 단계별 평가 기준 및 임계점에 관한 심도 있는 연구가 필요하다.

본 연구 결과를 바탕으로 다음과 같은 학술적, 정책적 시사점을 도출할 수 있다. 우선, 향후 침수를 포함한 자연재해 및 기후변화로부터 안전한 도시를 위해서는 피해액뿐만 아니라 인명피해를 함께 고려하는 것이 필요하다. 예를 들어, 현재 기후변화로부터 안전한 도시계획에 중요한 수단으로 활용되고 있는 방재지구 선정의 경우, “동일한 재해가 최근 10년 이내 2회 이상 발생하여 인명피해를 입은 지역”으로 선정 대상 조건에 피해액은 포함되어 있지 않다. 기후변화 도시재해 취약성 평가의 경우 민감도 평가에서 “최근 10년간 피해지역”을 반영하도록 되어 있는데 주로 침수지역 면적 및 위치가 분석에 활용된다. 한편, 앞서 논의된 것과 같이 침수 취약성에 대한 선행연구들의 경우 피해액에 초점을 맞추어 연구가 진행되어 왔다. 하지만, 본 연구에서 도출된 바와 같이 피해액과 인명피해에 영향을 미치는 요인이 상이하기 때문에 침수취약성에 대한 학술 연구 및 정책 수립에 있어 두 가지 모두에 대한 고려가 필요하다. 또한, 향후 침수피해액을 낮추기 위해서는 노후 단독주택 비율이 높고, 하천 및 수리시설 등이 위치하여 물리적 취약성이 높고, 재정 자립도는 낮은 지자체를 우선적으로 선별하여 국가 차원에서 방재정책을 지원할 필요가 있다. 특히, 피해액이 높은 지역은 물리적 취약요인이 높은 특성이 있으므로 노후 주거지 개선 및 하수관거 시설용량 증설, 도로의 투수성 포장재 교체 등 시설적 대응력 보강을 통해 피해액을 저감할 수 있을 것이다. 한편, 인명피해를 낮추기 위해서는 신체적·경제적 취약인구 거주 비율이 높고 시설적 대응력이 부족한 지역을 우선적으로 선별하여 지역의 재해 대응역량을 강화할 필요가 있다. 신체적 취약인구는 재해 발생 시 대응역량 부족으로 침수로 인한 사망 및 부상 사고가 발생할 위험이 높기 때문이다. 보건소는 어린이집이나 양로원, 병원 등 신체적 취약인구가 밀집한 시설에 방문하여 주기적으로 재해 대피요령 등 방재교육을 제공하고, 소방시설은 보육 시설과 요양 시설들의 위치 파악으로 재해 발생 시 우선 대피시설로 관리하는 방법 등을 통해 지역의 재해대응력을

강화할 수 있을 것으로 기대된다.

본 연구의 차별점은 침수피해 지표를 피해액과 인명피해로 구분하여 발생 양상을 분석하고, 침수피해와 침수취약요인 간 비선형적 관계를 기계학습 모형으로 밝혔다. 이는 피해액 영향요인과 인명피해 영향요인을 구분하고, 전통적인 회귀 모형에서 포착할 수 없었던 비선형적 관계를 밝힘으로써 기존 연구의 한계를 보완하였다. 한편, 본 연구의 한계는 다음과 같다. 재해연보에서 제공하는 데이터의 한계로 인해 피해액과 인명피해를 시·군·구 단위로 분석하였으며, 실수 또는 인재에 의한 인명피해를 구분하지 않았다. 향후 재해연보의 피해액과 인명피해 통계 데이터가 읍·면·동의 미시적 공간 단위에서 제공되고, 실족 및 인재에 의한 인명피해를 구분하여 제공된다면, 재해피해 발생에 영향을 미치는 지역의 공간적 특성을 면밀하게 식별할 수 있을 것이다. 후속연구로 자연재해 분야에서 사망 및 부상자에 대한 인명피해 환산 기준이 수립된다면, 보다 정확한 인명피해 강도를 반영할 수 있을 것이다. 한편, 도시지역과 비도시지역의 침수피해 차이에 대한 후속 연구는 도시지역과 비도시지역의 방재정책 및 계획 수립에 새로운 시사점을 제시할 수 있을 것이다.

주1. '철도사고 안전성능 평가(국토교통부, 2018)'에 따르면 철도사고 시 사망자 1을 부상자 100으로 환산하고, 「산업안전보건법 시행규칙」 제3조의 2 제1항 관련 [별표 1]에서 건설사업자 산업재해 발생 시 사망자 1에 대한 가중치는 부상재해자의 5배로 규정한다.

주2. 재해연보에서는 시·군·구 단위에서 재해원인별 피해액과 인명피해 통계자료를 공개하지 않고 있기 때문에, 본 연구에서는 시·군·구 자연재해 피해액과 인명피해 통계를 침수피해 자료로 활용하였다. 호우 및 태풍 피해는 2010년부터 2019년까지 10년간 우리나라 자연재해 피해액 중 89.42%, 인명피해의 65.94%를 차지하여 자연재해 피해의 주된 재해에 해당한다.

주3. 2017 인구주택총조사 전수집계 결과, https://kostat.go.kr/board.es?mid=a10301100200&bid=203&tag=&act=view&list_no=370326&ref_bid=203,236

주4. 한국국토정보공사, 침수 흔적도, <https://www.data.go.kr/data/15048628/fileData.do?recommendDataYn=Y> (2023년 3월부터 해당 공공데이터 개방 서비스가 중지됨)

주5. 서울연구원, 인포그래픽스, <https://www.si.re.kr/node/66289>

인용문헌 References

- 강정은·이명진, 2012. "퍼지모형과 GIS를 활용한 기후변화 홍수 취약성 평가: 서울시 사례를 중심으로", 『한국지리정보학회지』, 15(3): 119-136.
Kang, J.E. and Lee, M.J., 2012. "Assessment of Flood Vulnerability to Climate Change Using Fuzzy Model and GIS in Seoul", *Journal of the Korean Association of Geographic Information Studies*, 15(3): 119-136.
- 국립방재연구원, 2011. 「지역사회기반 홍수 피해저감지수 개발」, 국립방재연구원.
NDMI (National Disaster Management Institute), 2011. *Development of Community-based Flood Disaster Risk Reduction Index*, NDMI.
- 국토교통부, 2018. 「2016 철도사고 및 안전성능 분석보고서」, TS 교통안전공단.
MOLIT, 2018. *2016 Rail Accident and Safety Performance Analysis Report*, TS.
- 김도영·김광섭, 2018. "홍수피해유형별 홍수위험지수 적용성 분석", 『대한토목학회논문집』, 38(1): 29-39.
Kim, M. and Kim, G., 2018. "Analysis of the Applicability of Flood Risk Indices According to Flood Damage Types", *Journal of the Korean Society of Civil Engineers*, 38(1): 29-39.
- 김수진·서교·김상민·이경도·장민원, 2013. "침수피해 정보를 이용한 농경지의 지형학적 침수취약지도 작성: 진주시를 사례로", 『농촌계획』, 19(3): 51-59.
Kim, S.J., Seo, K., Kim, S.M., Lee, K.D., and Jang, M.W., 2013. "Mapping of Inundated Vulnerability Using Geomorphic Characteristics of Flood-damaged Farmlands: A Case of Jinju City", *Journal of Korean Society of Rural Planning*, 19(3): 51-59.
- 김지수·성효현·최광용, 2013. "도시 홍수 취약성의 공간적 분포: 서울 지역을 중심으로", 『한국지역지리학회지』, 19(4): 615-626.
Kim, J., Sung, H.H., and Choi, G., 2013. "Spatial Pattern of Urban Flood Vulnerability in Seoul", *Journal of the Korean Association of Regional Geographers*, 19(4): 615-626.
- 박준상·이수기, 2022. "해석가능한 기계학습을 활용한 보행만족 별 보행만족도 영향 요인 분석", 『국토계획』, 57(1): 26-41.
Park, J. and Lee, S., 2022. "Analysis of Influencing Factors of Walking Satisfaction by Purpose Using Interpretable Machine Learning", *Journal of Korea Planning Association*, 57(1): 26-41.
- 박한나·송재민, 2014. "침수 취약성 지표와 사회적 취약계층 비율 간의 상관관계 분석을 통한 환경정의 실증 연구", 『국토계획』, 49(7): 169-186.
Park, H. and Song, J., 2014. "Empirical Study on Environmental Justice through Correlation Analysis of the Flood Vulnerability Indicator and the Ratio of the Poor Population", *Journal of Korea Planning Association*, 49(7): 169-186.
- 박형준·서세교·이달별·정주철, 2020. "침수지역 내 노인 인구의 취약성에 관한 사례연구: 부산시 기장군을 중심으로", 『한국방재학회논문집』, 20(4): 33-41.
Park, H., Seo, S.G., Lee, D.B., and Jung, J., 2020. "A Case Study on the Vulnerability of the Elderly in Floodplains: Focused on Gijang, Busan", *Journal of the Korean Society of Hazard Mitigation*, 20(4): 33-41.
- 서만훈·이재송·최열, 2016. "공간적 자기상관성과 도시특성 요소를 고려한 자연재해 피해 분석", 『대한토목학회논문집』, 36(4): 723-733.
Seo, M.W., Lee, J.S., and Choi, Y., 2016. "Estimation of the Natural Damage Disaster Considering the Spatial Autocorrelation and Urban Characteristics", *Journal of the Korean Society of Civil Engineers*, 36(4): 723-733.

11. 서정석·한우석, 2019. “침수 취약지역과 사회적 취약계층의 공간적 상관성 분석: 제주특별자치도 사례를 중심으로”, 『한국방재학회논문집』, 19(4): 103-113.
Seo, J. and Han, W., 2019. “Exploring the Relationship Between Floods and Socially Vulnerable Groups: The Case of Jeju”, *Journal of the Korean Society of Hazard Mitigation*, 19(4): 103-113.
12. 신희재·최영제·이재웅, 2019. “지역적 특성을 고려한 도시지역 홍수 취약성 평가: 한강권역을 대상으로”, 『한국방재학회논문집』, 19(5): 293-303.
Shin, H., Choi, Y., and Yi, J., 2019. “Analysis of the Local Characteristics of Flood Damage Vulnerability in an Urban Area: The Han River Basin”, *Journal of the Korean Society of Hazard Mitigation*, 19(5): 293-303.
13. 심재현·김자은·이성호, 2012. “기후변화 대응을 위한 광역도시권 차원의 자연재해 저감방안 연구: 자연재해 취약성에 따른 수도권권의 유형화”, 『한국산학기술학회논문지』, 13(11): 5534-5541.
Shim, J.H., Kim, J.E., and Lee, S.H., 2012. “Classification of Cities in the Metropolitan Area based on Natural Hazard Vulnerability”, *Journal of the Korea Academia-Industrial Cooperation Society*, 13(11): 5534-5541.
14. 이경훈·류용욱·박종석, 2020. “2020년 8월 영산강섬진강 유역 홍수 피해조사 및 원인분석”, 『물과 미래』, 53(11): 21-48.
Lee, K.H., Ryu, Y.U., and Park, J.S., 2020. “Investigation and Cause Analysis of Flood Damage in the Yeongsan River Seomjin River Basin in August 2020”, *Water for Future*, 53(11): 21-48.
15. 이관형·이원호·양원직, 2016. “사회재난의 재난위해분석에 관한 연구”, 『한국방재안전학회논문집』, 9(2): 15-21.
Lee, K.H., Yi, W.H., and Yang, W.J., 2016. “A Study on Risk Analysis of Social Disaster”, *Journal of Korean Society of Disaster & Security*, 9(2): 15-21.
16. 이상혁·강정은·박창석, 2016. “베이지안 확률통계와 GIS를 연계한 기후변화 도시홍수 리스크 평가: 서울시 서초구를 대상으로”, 『한국지리정보학회지』, 19(4): 36-51.
Lee, S.H., Kang, J.E., and Park, C.S., 2016. “Urban Flood Risk Assessment Considering Climate Change using Bayesian Probability Statistics and GIS: A Case Study from Seocho-Gu, Seoul”, *Journal of Korean Association Geographic Information Studies*, 19(4): 36-51.
17. 이은석, 2020. “건축물 정보를 활용한 도시침수 취약성 진단방법 개발”, 『한국기후변화학회지』, 11(1): 65-75.
Lee, E., 2020. “Development of a Vulnerability Diagnosis Model for Urban Surface Water Floods using Building Big Data”, *Journal of Climate Change Research*, 11(1): 65-75.
18. 이제승·이현희, 2019. “기계학습 알고리즘을 이용한 보행만족도 예측모형 개발”, 『국토계획』, 54(3): 106-118.
Lee, J.S. and Lee, H., 2019. “Developing a Pedestrian Satisfaction Prediction Model Based on Machine Learning Algorithms”, *Journal of Korea Planning Association*, 54(3): 106-118.
19. 이종석·최현일, 2016. “연안지역 홍수피해자료 분석을 통한 홍수 취약성 평가”, 『한국방재학회논문집』, 16(2): 377-385.
Lee, J.S. and Choi, H.I., 2016. “Flood Vulnerability Assessment by Analysis of Flood Damage Data in Coastal Areas”, *Journal of Korean Society of Hazard Mitigation*, 16(2): 377-385.
20. 장미홍·박창열, 2019. “침수피해지역 토지이용 특성과 방재대책 연구: 제주지역 사례로”, 『한국방재학회논문집』, 19(7): 135-144.
Jang, M. and Park, C., 2019. “A Study on the Flood Protection Measures and Land Use Characteristics of Flood Damage Areas: The Case of the Jeju Area”, *Journal of the Korean Society of Hazard Mitigation*, 19(7): 135-144.
21. 조심정, 2018. “리스크 곡선을 활용한 사회재난 인명피해 위험도 분석 연구”, 『한국방재학회논문집』, 18(2): 93-103.
Cho, S., 2018. “A Study of Risk Analysis for Human Casualty of Social Disasters based on F-N Curve”, *Journal of the Korean Society of Hazard Mitigation*, 18(2): 93-103.
22. 조심정·최승용, 2018. “통계자료 활용 안전사고 위험도 분석을 통한 재난안전예산 중점투자 유형 선정”, 『한국방재학회논문집』, 18(4): 135-146.
Cho, S. and Choi, S., 2018. “Top Priority Accident Types in Disaster and Safety Budgets based Risk Analysis using Statistical Data”, *Journal of the Korean Society of Hazard Mitigation*, 18(4): 135-146.
23. 조현우·최승용, 2018. “피해 통계자료 활용 유클리드 거리 기반 재난 사고·유형간 위험순위 산출”, 『한국방재학회논문집』, 18(6): 133-147.
Cho, H. and Choi, S., 2018. “Risk Ranking among Types of Disaster and Safety Accidents Based on Euclidean Distance Using Damage Statistical Data”, *Journal of the Korean Society of Hazard Mitigation*, 18(6): 133-147.
24. 최열·서만훈, 2013. “도시특성 요소가 자연재해에 의한 피해액에 미치는 영향에 관한 연구”, 『국토계획』, 48(5): 309-318.
Choi, Y. and Seo, M.H., 2013. “Assessing the Impact of the Factors of Urban Characteristics on the Damages Caused by Natural Disaster Using Panel Model”, *Journal of Korea Planning Association*, 48(5): 309-318.
25. 최창현·김종성·김동현·이준형·김덕환·김형수, 2018. “머신러닝 기법을 이용한 수도권 지역의 호우피해 예측함수 개발”, 『한국방재학회논문집』, 18(7): 435-447.
Choi, C.H., Kim, J.S., Kim, D.H., Lee, J.H., Kim, D.K., and Kim, H.S., 2018. “Development of Heavy Rain Damage Prediction Functions in the Seoul Capital Area using Machine Learning Techniques”, *Journal of Korean Society of Hazard Mitigation*, 18(7): 435-447.
26. 최충익, 2003. “도시적 토지이용의 변화가 도시의 자연재해 취약성에 미치는 영향에 관한 연구: 경기도를 사례로”, 『국토계획』, 38(2): 35-48.
Choi, C.I., 2003. “A Study on Natural Hazards Vulnerability in Urban Area by Urban Land Use Change: In Case of Kyonggi Province”, *Journal of Korea Planning Association*, 38(2): 35-48.
27. 한국행정연구원, 2020. 『지방자치단체 재난안전예산관리 효과성 제고 방안 연구: 재난안전예산 사전검토제도 도입을 중심으로』, 한국행정연구원.
Korea Institute of Public Administration, 2020. *A Study on the Improvement of the Effectiveness of Local Government's Disaster Safety Budget Management: Focusing on the Introduction of the Preliminary Review System for Disaster Safety Budget*, Korea Insti-

- tute of Public Administration.
28. 행정안전부, 2020. 「2020 재해연보」, 행정안전부.
Ministry of the Interior and Safety, 2020. *The Annual Disaster Report in 2020*, Ministry of the Interior and Safety.
 29. 홍승진·김길호·김경탁, 2018. “자연재난 인명피해 현황 분석: 호우, 태풍을 중심으로”, 『물과 미래』, 51(4): 50-53.
Hong, S., Kim, G., and Kim, K., 2018. “Analyze Natural Disaster Casualties – with a Focus on Heavy Rains and Typhoons”, *Water for Future*, 51(4): 50-53.
 30. Alderman, K., Turner, L.R., and Tong, S., 2012. “Floods and Human Health: A Systematic Review”, *Environment International*, 47: 37-47.
 31. Ashley, S.T. and Ashley, W.S., 2008. “Flood Fatalities in the United States”, *Journal of Applied Meteorology and Climatology*, 47(3): 805-818.
 32. Bakkensen, L.A., Fox-Lent, C., Read, L.K., and Linkov, I., 2017. “Validating Resilience and Vulnerability Indices in the Context of Natural Disasters”, *Risk Analysis*, 37(5): 982-1004.
 33. Bigi, V., Comino, E., Fontana, M., Pezzoli, A., and Rosso, M., 2021. “Flood Vulnerability Analysis in Urban Context: A Socioeconomic Sub-Indicators Overview”, *Climate*, 9(1): 12.
 34. Chakraborty, L., Rus, H., Henstra, D., Thistlethwaite, J., and Scott, D., 2020. “A Place-based Socioeconomic Status Index: Measuring Social Vulnerability to Flood Hazards in the Context of Environmental Justice”, *International Journal of Disaster Risk Reduction*, 43: 101394.
 35. Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., and Bin Ahmad, B., 2020. “Modeling Flood Susceptibility using Data-driven Approaches of Naïve Bayes Tree, Alternating Decision Tree, and Random Forest Methods”, *Science of Total Environment*, 701: 134979.
 36. Chuang, M.T., Chen, T.L., and Lin, Z.H., 2020. “A Review of Resilient Practice based upon Flood Vulnerability in New Taipei City, Taiwan”, *International Journal of Disaster Risk Reduction*, 46: 101494.
 37. Cutter, S.L. and Finch, C., 2008. “Temporal and Spatial Changes in Social Vulnerability to Natural Hazards”, *PNAS*, 105(7): 2301-2306.
 38. Doocy, S., Daniels, A., Murray, S., and Kirsch, T.D., 2013. “The Human Impact of Floods: A Historical Review of Events 1980-2009 and Systematic Literature Review”, *PLoS Currents*, 5.
 39. Ha, J. and Kang, J.E., 2022. “Assessment of Flood-risk Areas using Random Forest Techniques: Busan Metropolitan City”, *Natural Hazards*, 111(3): 2407-2429.
 40. IPCC, 2007. *Climate Change 2007: Impacts, Adaptation and Vulnerability*, Cambridge, UK: Cambridge University Press.
 41. IPCC, 2014. *Climate Change 2014: Impacts, Adaptation, and Vulnerability*, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
 42. Jamshed, A., Birkmann, J., Feldmeyer, D., and Rana, I.A., 2020. “A Conceptual Framework to Understand the Dynamics of Rural-Urban Linkages for Rural Flood Vulnerability”, *Sustainability*, 12(7): 2894.
 43. Jeong, S.H. and Yoon D.K., 2018. “Examining Vulnerability Factors to Natural Disasters with a Spatial Autoregressive Model: The Case of South Korea”, *Sustainability*, 10(5): 1651.
 44. Jongman, B., Winsemius, H.C., Fraser, S.A., Muis, S., and Ward, P.J., 2018. “Assessment and Adaptation to Climate Change Related Floods Risks”, *Oxford Research Encyclopedia of Natural Hazard Science*, Oxford, UK: Oxford University Press.
 45. Jonkman, S.N. and Kelman, I., 2005. “An Analysis of the Causes and Circumstances of Flood Disaster Deaths”, *Disasters*, 29(1): 75-97.
 46. Jonkman, S.N., 2014. “Loss of Life Due to Floods: General Overview”, in *Drowning* edited by Bierens, J., 957-965, Berlin, Heidelberg: Springer.
 47. Kim, H., Marcouiller, D.W., and Woosnam, K.M., 2018. “Rescaling Social Dynamics in Climate Change: The Implications of Cumulative Exposure, Climate Justice, and Community Resilience”, *Geoforum*, 96(2018): 129-140.
 48. Lee, H.K., Bae, Y.H., Son, J.Y., and Hong, W.H., 2020. “Analysis of Flood-Vulnerable Areas for Disaster Planning Considering Demographic Changes in South Korea”, *Sustainability*, 12(11): 4727.
 49. Lowe, D., Ebi, K., and Forsberg, B., 2013. “Factors Increasing Vulnerability to Health Effects before, during and after Floods”, *International Journal of Environmental Research and Public Health*, 10(12): 7015-7067.
 50. Ma, M., Zhao, G., He, B., Li, Q., Dong, H., Wang, S., and Wang, Z., 2021. “XGBoost-based Method for Flash Flood Risk Assessment”, *Journal of Hydrology*, 598: 126382.
 51. Maantay, J. and Maroko, A., 2009. “Mapping Urban Risk: Flood Hazards, Race, & Environmental Justice in New York”, *Applied Geography*, 29(1): 111-124.
 52. Munyai, R.B., Musyoki, A., and Nethengwe, N.S., 2019. “An Assessment of Flood Vulnerability and Adaptation: A Case Study of Hamutsha Muungamunwe Village, Makhado Municipality”, *Journal of Disaster Risk Studies*, 11(2): a692.
 53. Nasiri, H., Mohd Yusof, M.J., and Mohammad Ali, T.A., 2016. “An Overview to Flood Vulnerability Assessment Methods”, *Sustainable Water Resources Management*, 2(3): 331-336.
 54. Parvin, F., Ali, S.A., Calka, B., Bielecka, E., Linh, N.T.T., and Pham, Q.B., 2022. “Urban Flood Vulnerability Assessment in a Densely Urbanized City using Multi-factor Analysis and Machine Learning Algorithms”, *Theoretical and Applied Climatology*, 149: 639-659.
 55. Pathak, S., Panta, H.K., Bhandari, T., and Paudel, K.P., 2020. “Flood Vulnerability and Its Influencing Factors”, *Natural Hazards*, 104(3): 2175-2196.
 56. Pham, B.T., Luu, C., Dao, D.V., Phong, T.V., Nguyen, H.D., Le, H.V., Von Meding, J., and Prakash, I., 2021. “Flood Risk Assessment using Deep Learning Integrated with Multi-criteria

- Decision Analysis”, *Knowledge-Based Systems*, 219: 106899.
57. Rahmati, O. and Pourghasemi, H.R., 2017. “Identification of Critical Flood Prone Areas in Data-scarce and Ungauged Regions: A Comparison of Three Data Mining Models”, *Water Resources Management*, 31(5): 1473-1487.
 58. Rufat, S., Tate, E., Emrich, C.T., and Antolini, F., 2019. “How Valid Are Social Vulnerability Models?”, *Annals of the American Association of Geographers*, 109(4): 1131-1153.
 59. Salas, J. and Yepes, V., 2018. “Urban Vulnerability Assessment: Advances from the Strategic Planning Outlook”, *Journal of Cleaner Production*, 179: 544-558.
 60. Salazar-Briones, C., Ruiz-Gibert, J.M., Lomelí-Banda, M.A., and Mungaray-Moctezuma, A., 2020. “An Integrated Urban Flood Vulnerability Index for Sustainable Planning in Arid Zones of Developing Countries”, *Water*, 12(2): 608.
 61. Sayers, P., Penning-Rowsell, E.C., and Horritt, M., 2018. “Flood Vulnerability, Risk, and Social Disadvantage: Current and Future Patterns in the UK”, *Regional Environmental Change*, 18: 339-352.
 62. Tascón-González, L., Ferrer-Julià, M., Ruiz, M., and García-Meléndez, E., 2020. “Social Vulnerability Assessment for Flood Risk Analysis”, *Water*, 12(2): 558.
 63. Tate, E., Rahman, M.A., Emrich, C.T., and Sampson, C.C., 2021. “Flood Exposure and Social Vulnerability in the United States”, *Natural Hazards*, 106: 435-457.
 64. Tehrany, M.S., Pradhan, B., and Jebur, M.N., 2014. “Flood Susceptibility Mapping using a Novel Ensemble Weights-of-evidence and Support Vector Machine Models in GIS”, *Journal of Hydrology*, 512: 332-343.
 65. Tellman, B, Schank, C., Schwarz, B., Howe, P.D., and de Sherbinin, A., 2020. “Using Disaster Outcomes to Validate Components of Social Vulnerability to Floods: Flood Deaths and Property Damage across the USA”, *Sustainability*, 12(15): 6006.
 66. Yoon, D.K. and Jeong, S., 2016. “Assessment of Community Vulnerability to Natural Disasters in Korea by Using GIS and Machine Learning Techniques”, in *Quantitative Regional Economic and Environmental Analysis for Sustainability in Korea* edited by Kim, E. and Kim B.H.S., 123-140, Springer Singapore.
 67. Yoon, D.K., 2012. “Assessment of Social Vulnerability to Natural Disasters: A Comparative Study”, *Natural Hazards*, 63(2): 823-843.
 68. Zahran, S., Brody, S.D., Peacock, W.G., Vedlitz, A., and Grover, H., 2008. “Social Vulnerability and the Natural and Built Environment: A Model of Flood Casualties in Texas”, *Disasters*, 32(4): 537-560.

Date Received	2022-11-30
Reviewed(1 st)	2023-01-02
Date Revised	2023-03-31
Reviewed(2 nd)	2023-04-07
Date Accepted	2023-04-07
Final Received	2023-05-22